Estimating Economic Consequences in Organizational Effectiveness Experiments

D. Epple, E. Fidler, P. Goodman

Graduate School of Industrial Administration
Carnegie-Mellon University
Pittsburgh, Pennsylvania 15213
Estimating Economic Consequences in Organizational Effectiveness Experiments

D. Epple, E. Fidler, P. Goodman

Graduate School of Industrial Administration Carnegie-Mellon University Pittsburgh, Pennsylvania 15213

This research was sponsored by:

1. the Organizational Effectiveness Research Program, Office of Naval Research (Code 452), under Contract No. N00014-75-C-0973; NR 170-801. Reproduction in whole or in part is permitted for any purpose of the United States Government. Approved for public release; distribution unlimited.

2. Under contract with the Institute for Social Research of The University of Michigan; the judgments expressed by the publisher do not necessarily represent the official opinion or policy of the Institute for Social Research or its affiliates.
<table>
<thead>
<tr>
<th>REPORT DOCUMENTATION PAGE</th>
<th>READ INSTRUCTIONS BEFORE COMPLETING FORM</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. REPORT NUMBER</td>
<td>4</td>
</tr>
<tr>
<td>2. GOVT ACCESSION NO.</td>
<td></td>
</tr>
<tr>
<td>4. TITLE (and Subtitle)</td>
<td>Estimating Economic Consequences in Organizational Effectiveness Experiments</td>
</tr>
<tr>
<td>7. AUTHOR(s)</td>
<td>D. Epple, E. Fidler, N. Goodman</td>
</tr>
<tr>
<td>9. PERFORMING ORGANIZATION NAME AND ADDRESS</td>
<td>Graduate School of Industrial Administration Carnegie-Mellon University Pittsburgh, Pa. 15213</td>
</tr>
<tr>
<td>10. PROGRAM ELEMENT, PROJECT, TASK AREA &amp; WORK UNIT NUMBERS</td>
<td>NR 170-801</td>
</tr>
<tr>
<td>11. CONTROLLING OFFICE NAME AND ADDRESS</td>
<td></td>
</tr>
<tr>
<td>13. NUMBER OF PAGES</td>
<td>3/31/77</td>
</tr>
<tr>
<td>14. MONITORING AGENCY NAME &amp; ADDRESS (IF DIFFERENT FROM CONTROLLING OFFICE)</td>
<td>Organizational Effectiveness Research Programs Office of Naval Research (Code 452) Arlington, Va. 22217</td>
</tr>
<tr>
<td>15. SECURITY CLASS. (OF THIS REPORT)</td>
<td>Unclassified</td>
</tr>
<tr>
<td>17. DISTRIBUTION STATEMENT (OF THE ABSTRACT ENTERED IN BLOCK 20, IF DIFFERENT FROM REPORT)</td>
<td></td>
</tr>
<tr>
<td>18. SUPPLEMENTARY NOTES</td>
<td></td>
</tr>
<tr>
<td>19. KEYWORDS (CONTINUE ON REVERSE SIDE IF NECESSARY AND IDENTIFY BY BLOCK NUMBER)</td>
<td>Organizational effectiveness, productivity, assessing productivity, multivariate models</td>
</tr>
<tr>
<td>20. ABSTRACT (CONTINUE ON REVERSE SIDE IF NECESSARY AND IDENTIFY BY BLOCK NUMBER)</td>
<td>This report presents three models for assessing one measure of organizational effectiveness--productivity. The models are delineated conceptually, empirically tested, and their advantages and disadvantages reviewed.</td>
</tr>
</tbody>
</table>

This report presents three models for assessing one measure of organizational effectiveness--productivity. The models are delineated conceptually, empirically tested, and their advantages and disadvantages reviewed.
Estimating Economic Consequences in
Organizational Effectiveness Experiments

1. The Problem

A major thrust in current research on organizations is experimentation with new forms of work organization. The purpose of this research is to increase organizational effectiveness in both the public and private sectors. Two dimensions of organizational effectiveness which have received most of the attention are productivity and psychological outcomes from work. This paper examines methods to evaluate changes in productivity from an organizational intervention.

Most new forms of work organization experiments assume that changes can be brought about in economic dimensions such as productivity. Increases in productivity are beneficial to management and lead to greater earning opportunities for the worker. The problem posed by this assumption is how to determine whether productivity has increased.

A cursory review of the current literature on organizational effectiveness will show that there is no substantial evidence to document changes in productivity. Clearly there are a lot of claims. Both consultants and managers have said that productivity in the experimental organizations has increased (cf. Glaser, 1975); however, the validity of this information is questionable given the propensity of people involved in a study to overestimate the results (Gordon, 1975).

"Substantial" evidence means that quality data is available over a
sufficiently long period of time and that there are sufficient controls available to separate out the effects of the experimental intervention from other variables that influence productivity. Most studies (cf. Glaser, 1975) simply assert that productivity is improved or increased by some percent without detailing the source of the data or identifying the cause of the change. The basic thesis in this paper is that production is a function of a number of controllable and uncontrollable variables. Changes in production can be affected by any combination of these variables. An organizational intervention simply adds another variable to this complex production function. Separating its effect is a complex analytical task.

Why have we not made more progress in estimating productivity differences? One reason is that much of the research in this area has been done by social scientists who are more interested in psychological than economic outcomes of work. Another reason is that many people involved in introducing organizational change have been more concerned with the process of change than with the outcomes. In other cases companies have not been willing to make economic data available, or the resources have not been available to analyze economic data. Another reason, and the central rationale for this paper, is that there have not been good analytic models available to assess productivity changes in organizational experiments. That is, we do not have a good methodology for estimating productivity changes.
2. The Setting

This paper examines several models used to estimate productivity differences in an experimental intervention in a coal mine. The forms of the models are presented at a general level to enable the models to be generalized to other settings. Clearly, some of the variables in the models will change in different settings, but the form of each model and the testing procedures will not.

The Rushton Mining Company entered into a Quality of Work (QOW) experiment in 1973 (Goodman 1977). One of the three sections (mining departments) became an experimental section. Autonomous work group teams were created in that section by outside consultants. While it is not possible to detail the experimental intervention (see Goodman, 1976), the communication, decision making, authority, and reward systems of that section were substantially modified. The basic hypothesis was that these changes would increase productivity levels and the quality of working life.

The design of this study permits comparing productivity of the experimental section against two other mining sections. In a sense these sections can serve as control groups to assist in isolating the effect of the experimental intervention on productivity. However, since these three sections are not perfectly matched as to men, machinery, and physical conditions the design is not truly an experimental design. At best this study can be classified as a quasi-experimental design; contrasts within the experimental section over time or between the experimental and control sections cannot
definitively isolate the experimental effect on productivity.

A coal mine's production is generally stated in terms of tons produced (per section). The set of variables affecting production is complex. Some variables are controllable, such as the number of men working or the type of equipment. Other variables are uncontrollable such as the character of roof and runway conditions. In comparing differences in tons produced per section it is important to understand which variables contribute to production. For example, if the experimental section outproduces the control sections but the latter have very poor roof conditions, then the difference in productivity may be caused by uncontrollable physical conditions or the experiment. To analyze productivity differences we need to describe a production function which includes the major predictor variables, and then to identify whether there are shifts in the coefficients of these variables over time.

3. Testing Procedures

Since it is not possible to control nonexperimental variables (e.g., roof conditions) which differentially affect the experimental and control groups, it is necessary to statistically control for these differences. We have adopted three alternative methods for testing for productivity changes after correcting for changes in uncontrollable variables across the experimental and control sections. (Where not indicated otherwise, the experimental section prior to the experiment
and the nonexperimental sections will be referred to as the control sections).

All three methods seek to identify differences in the structure of the production function (i.e., the relationship between resources used and quantities produced) between the experimental section and the control sections. Thus we began by estimating a separate production function for each section for each year. The appropriate positive or negative signs can be specified a priori for almost all variables in the production function equation. For example, output should be positively related to working time and to the roof and runway condition variables. Before accepting the estimated equation for an individual section, we have required that all coefficients which are significantly different from zero have the predicted sign. In the few instances in which initial estimates did not satisfy the requirement, we have attempted to determine whether exceptional circumstances (e.g., unusual physical conditions) in a particular section might have given rise to an unreasonable coefficient estimate. In some cases, dummy variables were introduced to account for such special circumstances.

The first and simplest procedure is to test for differences in the constant term in the production function across sections. Differences, if any, in the sections, including differences caused by the experiment, are assumed to affect only the constant term. Given this assumption, the homogeneity of the sections in the control
group or a subset of the control group can be tested by an analysis of variance, i.e., by testing whether a significantly higher proportion of variance is explained by allowing a different constant term for each section than by imposing the same constant term across all sections. If the experiment enhanced productivity, this would be reflected by a higher constant term in the production function for the experimental section than in the production functions for the control sections. The t-test can be applied to determine whether the difference in the constant term between the experimental and control groups is significant.

The above approach has the advantage of being both simple to apply and easy to interpret. The disadvantage of the procedure is the assumption that the coefficients on both the controllable and uncontrollable variables are the same across the experimental and control sections. It is not obvious, a priori, that this should be the case. For example, if the result of the experiment is that the men work more efficiently in the presence of adverse conditions, this would be reflected on the coefficients of the conditions variables rather than the intercept. The two additional tests described below do not require such restrictive assumptions.

A second method is to test for differences in any of the estimated coefficients of the production function across the sections. The maintained (null) hypothesis is that all the coefficients in the production function are the same in all sections. The alternative hypothesis is that one or more of the coefficients differ across
sections. Thus, if the experiment resulted in a significant change in
the production function in the experimental section, the maintained
hypothesis should be rejected when the experimental section is
compared to the control sections.

The maintained hypothesis in the first step of the test is that
the coefficients in all six sections are the same. The alternative
is that the coefficients in all six are different. Since the tests
are sequential, the form of each subsequent test will depend on the
outcome of the one preceding.1 One simple sequence is as follows:
If the null hypothesis of the first test is rejected, the production
functions of the sections are not all alike. The second step would
then be to test whether the production function is the same for all
control sections. If so, the control groups would be combined to
yield a single production function, and this function would then be
compared to the experimental section. Selective testing of subsets
of the coefficients would determine more precisely which of the co-
eficients differ between the experimental and control section pro-
duction functions.

A great many other outcomes are possible. The control sections
might be found to differ among themselves. The sections might be the

1 The probability of obtaining a given conclusion is not in-
dependent of the test sequence. Ideally, we would like to compute the
probability of obtaining a particular conclusion taking account of the
test sequence. However, in the absence of knowledge of the true values
of the parameters of the model, such a calculation is not possible. Thus,
confidence regions for a given level of significance are determined
by treating each step of the test sequence as independent of the outcomes
of previous steps in the sequence. This rather unsatisfying procedure
appears to be widely employed since no generally applicable alternative
procedure is available.
same in each year and different across years. The control and experimental sections might be found to be all alike. Rather than enumerate all possible sequences, we will defer further discussion until the results section where the actual test sequence is presented.

The second testing procedure described above also has several limitations. The power of the analysis of variance test is much greater when one uses a single production function for the five control sections in a test against the experimental section than when one tests the five control sections, each with a different production function, against the experimental section. For a given level of significance, one may get ambiguous results from the analysis of variance test because the power of the test changes when different combinations of coefficients or sections are tested. A second problem is that individual coefficients may be different in the experimental section relative to the control sections, but some may be higher and others lower so that the net effect on productivity would not be clear cut.

The third method used to test the experimental section against the control sections is designed to provide an alternative which would potentially be conclusive if the analysis of variance test were

---

2 Here each set of annual data is counted separately since a separate production function is estimated for each section for each year.
ambiguous for either of the reasons identified above. This procedure requires only the production functions for the control sections. The values of the variables observed in the experimental section are substituted into the estimated production functions for the control sections. This provides a prediction of the amount which would have been produced had the resources from the experimental section been used in the control sections under similar physical conditions. The actual average weekly production from the experimental section can then be compared to the predicted average from the equations for the control sections. If the predicted amount from the control sections was significantly lower (higher) than actual production from the experimental section, one would conclude that under similar physical conditions the experimental section was more (less) productive than the control sections. The relevant test statistic has the \( t \) distribution. The derivation of the test statistic is provided in the Appendix.

Relative to the second method, the third method of comparison has the advantage of simplicity; the information on productivity is summarized in a single statistic comparing mean actual output of the experimental section with mean predicted output from using the same resources in the control section. The second advantage relative to both of the alternative tests is that one need not estimate a production function for the experimental section. For purposes of this test, any shifting of the production function of the experimental
section will be reflected in the production of that section. Finally, the power of the test is greater than the analysis of variance test because the latter test compares the sections along several dimensions (as many dimensions as there are estimated coefficients) while the former test is based on a single dimension (production).

The evaluation results are not based exclusively on any one of the above tests. The multidimensional nature of the analysis of variance test enables one to identify differences among the sections which may not be reflected in the production figures. These differences are of interest in themselves as a supplement to the simpler test of differences in productivity. The test for differences in the constant terms is the logical first step in attempting to pinpoint differences in the production functions.

4. Methodology

Sample

The analyzed data set consists of 92 weeks in 1973 and 1974 for the three sections—experimental (South), North, and East. The sample year 1973 starts at 1/2/73 and ends at 12/1/73. From these 48 weeks two were deleted which represent the miners' vacation period in July. Therefore, the 1973 data are based on 46 weeks.

The experimental year begins in the first week of December 1973 and runs through 11/9/74. From these 49 weeks the two-week vacation plus another week lost from the memorial week strike (8/17/74 to 8/24/74) were deleted leaving 46 weeks for analysis. In the remaining period in November and in the first week of December there was no production because of the national coal strike. Since there
were only a few remaining actual production weeks in December and since the work in that period was under a new contract, we decided to include those weeks in the 1975 data.

Variables

Table 1 describes the major variables used in the analysis. The label or acronym, variable name, and description are given; then the operational form and source of the data are identified. The means and standard deviations are for all three sections for the 1973 and 1974 time periods.

Our prior expectations about the coefficients are indicated in Table 2. Since we are using a linear approximation to the production function, and the conditions variables may shift the function up or down, we do not have a priori expectations concerning the sign of the constant term. Improvements in conditions should increase production as indicated by the positive signs on the physical conditions variables in the table. The moves variable is designed to measure the distance of the miner from the feeder and should thus be negative. Since pillaring is more productive than developmental mining we expected a positive effect on the pillaring variable. Increases in crew size should enhance production giving a positive sign to the man-days variable. Delays were expected to have no effect beyond the reduction in working time. Since delays have been deducted from potential crew time in constructing ACWOT, we expect zero coefficients
on the delay variables when delays are entered separately. Crew time should have a positive effect on production.

5. Results

The production function estimates are presented in Table 3. A comparison of the coefficient estimates across the various sections reveals that, where coefficients are statistically significant, they do have the anticipated sign.

We now turn to the results derived from our tests. The first method of testing will serve to indicate whether there were differences in any of the intercepts of the model across the sections.

In row 1 of Table 4 we require the intercept for the five control sections to be the same and test to determine whether the intercept for South 74 (experimental) is significantly different. The estimated difference of -48.9 is not significant. In row 2, all of the section intercepts are allowed to be different with South 73 used as a reference. The coefficients for North 73 and North 74 are significantly lower than the coefficients for South 73 at the five percent level, and the coefficient for East 74 is significantly lower at the ten percent level. This result indicates that the intercepts for the control group are different, and this is confirmed by row 3 of Table 4 which includes only the five control sections. North has a significantly lower intercept than the remaining control sections in both years. In the remaining three rows, the intercept of each
section in 1974 is compared to the 1973 intercept. The difference is significant only in North.

The results in Table 4 provide no evidence that the intercept has been shifted by the experiment. This conclusion emerges in the first row when South 74 is compared to the entire control group and in the last row when South 74 is compared to South 73. The results in Table 4 do, however, suggest that there are significant differences among the intercepts of the sections in the control group. The results of method two below provide a more general test for differences in any of the coefficients across the sections.

In row 1 of Table 5 we test whether individual production functions for the six sections fit the data significantly better than a single function applied to all six. At the five percent significance level, the sections are not significantly different. Since the power of the F-test varies considerably when various sections are combined, we present the results of several alternative tests. The results in the second row indicate that the control sections are not significantly different. When a single function for the control section is tested against the experimental section in row 3, the difference proves to be significant at the one percent level.

Further evidence of the similarity of the control sections is provided in rows 4 and 5. In row 4, a single function for the three sections in 1973 is tested against separate functions for each section and no significant difference is found. In row 5, the combined control
group in 1973 is compared to the combined control group in 1974
and again no difference is found.

Row 6 indicates no significant difference across experimental
and control groups in 1974. However, when the experimental and
control groups are combined in 1974 and tested against the combined
control group in 1973, a significant difference is found as in-
dicated in row 7. This should be contrasted to the results in 5
where no significant differences were found when the experimental
section was not included in 1974. Finally, the models for each of
the sections are compared across years in rows 8, 9, and 10. Here
it is found that one of the control sections differs between 1973 and
1974, but the production function for the experimental section in 1974
is not significantly different from the production function in 1973.

Additional tests were conducted to explore possible differences
between South 74 and the control group. In these tests additional
dummy variables were included to allow the ROOF, RUNWAY, and PILLARINC
coefficients to differ for South 74 when it was included in estimating
a single production function for the experimental and control groups
combined. After differences are allowed in the coefficients of the
three variables identified above, the production functions for the
experimental and control sections are not significantly different.

The conclusion of the F-test is that there is some relatively
weak evidence that the experimental and control sections are different.
The differences are attributable to the ROOF, RUNWAY, and PILLARINC,
coefficients. These differences in the coefficients of the conditions variables may be attributable to the experiment though such differences would not have been predicted a priori. For example, it may be that the experiment caused the crew to work at the same productivity level independent of runway conditions giving rise to the insignificant coefficient on runway in Table 3 for South 74. The higher coefficient on pillaring indicated greater productivity in pillaring in the experimental section. If the higher coefficient on roof is to be attributed to the experiment, one would have to conclude that the experiment made productivity more sensitive to changes in roof conditions. Since the coefficients on roof and runway in South 74 differ in opposite directions from those of the control group, sampling errors rather than experimental effects may be the cause of these differences.

The results of the test of intercepts appear in some respects inconsistent with the results of the analysis of variance test. The intercept tests indicated no significant effect of the experiment while the analysis of variance tests indicate a significant difference between the experimental and control groups. Also, the intercept tests indicated that North differed significantly from the other control sections while the analysis of variance tests suggest that what differences exist are attributable to East.

The explanation of these seemingly inconsistent results is traceable to the underlying assumptions and the relative power of the tests. The intercept test is predicated on the assumption that differences, if any, will be reflected in the intercepts. Offsetting differences in the
other coefficients may not be picked up by the intercept test. In contrast, the analysis of variance test allows for quite general differences, but the discriminatory power of the test is lower than that of the intercept test. Clearly, it is desirable to have a single summary measure of the effects of the experiment. The intercept test is unsatisfactory because there is no a priori reason to expect only the intercept to be affected. The analysis of variance test identifies differences but does not indicate the net effect of those differences on productivity. It is for these reasons that we developed the third method of testing for differences in productivity.

Both the intercept test and the analysis of variance test cast doubt on the hypothesis that all sections in the control group have the same production function. Therefore, we conducted the means test not only with the combined control group, but also using only South 73 as the control group. If unobserved variables are responsible for differences among the control sections, South 73 may be a more reliable control since such unobserved variables may differ less over time for a given section than they vary across sections.

The results of means tests are presented in Table 6. For the results in the upper half of the Table, the production function for South 73 was used to obtain predicted average weekly production. This prediction is obtained by substituting the independent variables observed in each week in South 74 into the estimated production function for South 73. The average of these predictions, denoted \( y_p \),
was then subtracted from actual average production in South 74 to obtain the estimated difference, d. The difference is an estimate of the amount by which average weekly production in South 74 exceeded the amount which would have been produced had the same resources been used under similar physical conditions in South 73. The results in the bottom half of the Table were obtained by applying the same procedure using the production function estimated when the data from all the control sections were combined.

The results indicate that average tonnage produced per week in South 74 was not significantly different from the amount which would have been produced had the same resources been used in the nonexperimental sections under similar physical conditions. This conclusion is obtained using either South 73 or the combined nonexperimental sections as the control.

6. Discussion-Summary

The general results do not indicate any increase in productivity in the experimental section during the first experimental year. The reasons for these results can be classified as methodological or conceptual. One methodological problem may be that we have not stated the models correctly. That is, variables might be missing or some of the selected variables incorrectly operationalized. In reviewing the models we have used in this analysis there are some coefficients that appear different from what we would have predicted.
and in some cases we have introduced dummy variables to deal with special problems (e.g., special physical conditions) not captured by our regular measures. To deal with this issue we thought carefully about the appropriate conceptual model. We did not let either techniques or the data primarily guide our analysis. When anomalies have appeared in the data we have gone back to the original data sources to identify the actual daily events that were going on. To avoid ex post solution from this more detailed analysis, we always tried to recast these findings back into our conceptual framework.

Another methodological issue concerns the control groups. The lack of improvement in productivity might be explained by the fact that the other sections are not equivalent to the experimental section and the comparison with these groups is inappropriate. We have tried to respond to this question by developing a model that will be generalizable across sections. Indeed, the analysis examining the production functions across the different sections shows a great deal of similarity. The other analytic approach has been to treat each section as its own control group. That is, we compared the performance of the experimental section in the baseline with its performance in the experimental year. The overall strategy then is to acknowledge that there is a problem of equivalence, but at the same time to use analysis procedures and to consider alternative control groups. If we get a consistent picture across these different strategies we can feel more certain about the validity of the results.
A third issue concerns problems of measurement. Unreliability in either the independent or dependent variables clearly will confound the analysis. To some extent this problem is outside of our hands because we do not have the resources to set up our own measurement system for the economic variables. In some cases we have analytically examined some of the ordinal scales used for measuring physical conditions to assess the validity of those measures. For the dependent variable we intend to relate the current figure against other measures of the same phenomena during the coming year to get a measure of concurrent validity.

Another measurement issue concerns potential biases in the measures. Here we would find a high degree of consistency (reliability) but validity would be low. The issue is that certain bosses could have overstated production as a reaction to the experimental induction. Bosses in the experimental section might have wanted to make the section look better while bosses in the control groups might have overstated production to make their section look comparable to the experimental section. Our own observations of this problem are the following: first, if overstating production figures by the foremen occurred it was generally a reaction to higher level management's demands for greater production. Our information suggests this pressure was in evidence during the baseline and experimental years and is a constant across sections. Second, our observations and interviews with the foremen in the experimental section indicated that overstating
production did not occur as a function of the experiment; it might have occurred from production pressure. Third, there was some evidence that in some of the control crews production may have been overstated as a reaction to the experiment. This behavior however was not widespread and should not have affected the comparisons when we treated the experimental section as its own control.

The fourth issue concerns the level of aggregation used in this analysis. Our data have been examined on a weekly basis. It could be argued that our estimates would be more refined if we moved to a daily basis. This is an issue we hope to explore with the data from 1975 and 1976.

The last issue concerns the time period for analysis and this might be the most compelling issue. It takes a long time period to identify the results of a major change effort like the Rushton project. We are only evaluating performance during the first year. It is clearly appropriate to withhold any conclusions about productivity until the 1975 and 1976 data is examined.

7. Conclusions

In selecting a procedure for evaluating organizational effectiveness experiments, the following criteria are applicable. First, the procedure should be relevant, that is, it should offer the potential of answering questions of interest concerning the effects of the experiment. Second, the conditions required for applicability of the
procedure should be satisfied by the experimental environment.

In evaluating the economic consequences of an experiment, the primary question of interest will be determining whether the experiment increased productivity. If a change in productivity occurred, one may also want to pinpoint how the production function differs after the introduction of the experiment. Even if no effect on productivity is realized, there may be some interest in testing whether the experiment had any impact at all on the production process.

Of the three procedures considered in this paper, the two most appropriate for testing for differences in productivity are the intercept test and the means test. While both tests satisfy the first criterion, the means test will generally be preferable by the second criterion. The intercept test is based on the assumption that the experiment affects only the intercept and no other coefficients in the production function. There was no a priori reason for expecting this to be the case in the experiment discussed in this paper, and it is doubtful that one would generally have such strong prior information about the way in which an experiment might potentially affect the production function. The means test is particularly suitable for experiments in which the effects on the production function are uncertain since the test requires no assumption about the way in which the production function is changed.

Whether or not a change in productivity occurs, one may wish to test whether the production function is changed at all by the
experiment. In addressing this question, the analysis of variance test is appropriate since it is intended to identify differences in any or all coefficients between the experimental and control groups. Should evidence of differences be found, tests for differences in subsets of the coefficients can be applied to pinpoint the differences.
<table>
<thead>
<tr>
<th>ACRONYM</th>
<th>Variable Name</th>
<th>Meaning of Variable</th>
<th>Operational Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>COALBO</td>
<td>Coal and boney height added together</td>
<td>Height from ceiling to the floor</td>
<td>Inches</td>
</tr>
<tr>
<td>ROOF</td>
<td>Roof conditions</td>
<td>Quality of roof</td>
<td>1 to 5 scale</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1 very bad 5 very good</td>
</tr>
<tr>
<td>RUNWAY</td>
<td>Runway conditions</td>
<td>Quality of runway</td>
<td>1 to 5 scale</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1 very bad 5 very good</td>
</tr>
<tr>
<td>PILLARING</td>
<td>Pillaring</td>
<td>(See Baseline Report, 1975)</td>
<td>Number of shifts pillared per week</td>
</tr>
<tr>
<td>ABSEN</td>
<td>Number of total absences</td>
<td>Number of total absences</td>
<td>Total absences per week</td>
</tr>
<tr>
<td>MOVES</td>
<td>Major moves</td>
<td>Measure of the distance between face and feeder</td>
<td>Assumes the value zero for the weeks with a major move (i.e., a move of more than 299 minutes) and increases by one for each week without a major move</td>
</tr>
<tr>
<td>AVMAN</td>
<td>Average man days</td>
<td>Average man days worked per day of the week</td>
<td>Total man days worked per week</td>
</tr>
<tr>
<td>MOMINDEL</td>
<td>Moves and miner delays</td>
<td>Moves and miner delays</td>
<td>Moves and miner delays minutes</td>
</tr>
<tr>
<td>COMDEL</td>
<td>Combined delays</td>
<td>Combined delays</td>
<td>Car, machinery, bolter, physical, and miscellaneous delays-minutes</td>
</tr>
<tr>
<td>EXOUTDEL</td>
<td>Autonomous work group obligation and outside delays</td>
<td>Autonomous work group and outside delays</td>
<td>Autonomous work group obligation and outside delays minutes</td>
</tr>
<tr>
<td>ACWOT</td>
<td>Actual crew working time</td>
<td>Actual crew working time</td>
<td>Maximum possible crew time (=5850)-total delays minutes</td>
</tr>
<tr>
<td>DU</td>
<td>Denotes dummy variable</td>
<td>Introduced for periods of abnormal conditions</td>
<td>One during weeks when abnormal conditions prevail</td>
</tr>
<tr>
<td>Operational Form</td>
<td>Source</td>
<td>Mean</td>
<td>Standard Deviation</td>
</tr>
<tr>
<td>----------------------------------------</td>
<td>---------------------------------------------</td>
<td>------</td>
<td>--------------------</td>
</tr>
<tr>
<td>Inches</td>
<td>Company Records</td>
<td>64.0</td>
<td>4.8</td>
</tr>
<tr>
<td>1 to 5 scale</td>
<td>Rating from superintendent</td>
<td>3.9</td>
<td>1.3</td>
</tr>
<tr>
<td>1 very bad 5 very good</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 to 5 scale</td>
<td>Rating from superintendent</td>
<td>3.5</td>
<td>1.7</td>
</tr>
<tr>
<td>1 very bad 5 very good</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of shifts pillarating per week</td>
<td>Company records</td>
<td>1.13</td>
<td>3.4</td>
</tr>
<tr>
<td>Total absences per week</td>
<td>Company records</td>
<td>4.8</td>
<td>4.0</td>
</tr>
<tr>
<td>Assumes the value zero for the weeks</td>
<td>Created by analysts based on moves delay</td>
<td>1.7</td>
<td>1.8</td>
</tr>
<tr>
<td>with a major move (i.e., a move of</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>more than 299 minutes) and increases</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>by one for each week without a major</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>move</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total man days worked per week</td>
<td>Created by analysts based on company records</td>
<td>18.4</td>
<td>2.6</td>
</tr>
<tr>
<td>days per week worked</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aoves and minor delays-minutes</td>
<td>Created by analysts based on company records</td>
<td>452.9</td>
<td>428.0</td>
</tr>
<tr>
<td>Car, machinery, bolter physical,</td>
<td>Created by analysts based on company records</td>
<td>782.2</td>
<td>650.6</td>
</tr>
<tr>
<td>and miscellaneous delays-minutes</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Autonomous work group obligation and</td>
<td>Created by analysts based on company records</td>
<td>79.3</td>
<td>275.9</td>
</tr>
<tr>
<td>outside delays-minutes</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Maximum possible crew working time</td>
<td>Created by analysts based on company records</td>
<td>4231.0</td>
<td>894.0</td>
</tr>
<tr>
<td>(=5850)-total delays minutes</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>One during weeks when abnormal</td>
<td>Created by analysts based on company records</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>conditions prevail</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Independent Variables</td>
<td>Anticipated Sign</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-----------------------</td>
<td>------------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CONSTANT</td>
<td>?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CCALBO</td>
<td>+</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ROOF</td>
<td>+</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RUNWAY</td>
<td>+</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ABSEN</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MOVES</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PILLARING</td>
<td>+</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AVTRAN</td>
<td>+</td>
<td></td>
<td></td>
</tr>
<tr>
<td>M0M:INDEL</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>COM:DEL</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AGWOT</td>
<td>+</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### Table 3

**Production Model II**

**Medium Aggregated Delays**

<table>
<thead>
<tr>
<th>East 73</th>
<th>North 73</th>
<th>South 73</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Coefficient</strong></td>
<td><strong>T Value</strong></td>
<td><strong>Coefficient</strong></td>
</tr>
<tr>
<td>CONSTANT</td>
<td>1845.</td>
<td>0.8051</td>
</tr>
<tr>
<td>COALBO</td>
<td>-49.33</td>
<td>-1.514</td>
</tr>
<tr>
<td>ROOF</td>
<td>-71.30</td>
<td>-0.9155</td>
</tr>
<tr>
<td>RUNWAY</td>
<td>178.7</td>
<td>3.268</td>
</tr>
<tr>
<td>ABSEN</td>
<td>-12.47</td>
<td>-0.5925</td>
</tr>
<tr>
<td>MOVES</td>
<td>-194.7</td>
<td>-2.404</td>
</tr>
<tr>
<td>AVMAN</td>
<td>45.21</td>
<td>1.472</td>
</tr>
<tr>
<td>EXOUTDEL</td>
<td>-1294.</td>
<td>-1.168</td>
</tr>
<tr>
<td>MOMINDEL</td>
<td>-0.1949</td>
<td>-0.7866</td>
</tr>
<tr>
<td>COMDEL</td>
<td>-0.05737</td>
<td>-0.3810</td>
</tr>
<tr>
<td>ACWOT</td>
<td>0.729</td>
<td>11.35</td>
</tr>
</tbody>
</table>

---

<table>
<thead>
<tr>
<th>East 74</th>
<th>North 74</th>
<th>South 74</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Coefficient</strong></td>
<td><strong>T Value</strong></td>
<td><strong>Coefficient</strong></td>
</tr>
<tr>
<td>CONSTANT</td>
<td>-2338.</td>
<td>-1.660</td>
</tr>
<tr>
<td>COALBO</td>
<td>-22.32</td>
<td>-0.8761</td>
</tr>
<tr>
<td>ROOF</td>
<td>389.5</td>
<td>2.987</td>
</tr>
<tr>
<td>RUNWAY</td>
<td>11.60</td>
<td>0.1750</td>
</tr>
<tr>
<td>PILLARING</td>
<td>22.11</td>
<td>2.167</td>
</tr>
<tr>
<td>ABSEN</td>
<td>30.73</td>
<td>1.888</td>
</tr>
<tr>
<td>MOVES</td>
<td>-32.67</td>
<td>-1.190</td>
</tr>
<tr>
<td>AVMAN</td>
<td>196.6</td>
<td>3.008</td>
</tr>
<tr>
<td>EXOUTDEL</td>
<td>-0.3979</td>
<td>-0.4405</td>
</tr>
<tr>
<td>MOMINDEL</td>
<td>-0.2954</td>
<td>-1.667</td>
</tr>
<tr>
<td>COMDEL</td>
<td>-0.0186</td>
<td>-0.1123</td>
</tr>
<tr>
<td>ACWOT</td>
<td>1.000</td>
<td>10.41</td>
</tr>
</tbody>
</table>

---

<table>
<thead>
<tr>
<th>East 74</th>
<th>North 74</th>
<th>South 74</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Coefficient</strong></td>
<td><strong>T Value</strong></td>
<td><strong>Coefficient</strong></td>
</tr>
<tr>
<td>CONSTANT</td>
<td>-2338.</td>
<td>-1.660</td>
</tr>
<tr>
<td>COALBO</td>
<td>-22.32</td>
<td>-0.8761</td>
</tr>
<tr>
<td>ROOF</td>
<td>389.5</td>
<td>2.987</td>
</tr>
<tr>
<td>RUNWAY</td>
<td>11.60</td>
<td>0.1750</td>
</tr>
<tr>
<td>PILLARING</td>
<td>22.11</td>
<td>2.167</td>
</tr>
<tr>
<td>ABSEN</td>
<td>30.73</td>
<td>1.888</td>
</tr>
<tr>
<td>MOVES</td>
<td>-32.67</td>
<td>-1.190</td>
</tr>
<tr>
<td>AVMAN</td>
<td>196.6</td>
<td>3.008</td>
</tr>
<tr>
<td>EXOUTDEL</td>
<td>-0.3979</td>
<td>-0.4405</td>
</tr>
<tr>
<td>MOMINDEL</td>
<td>-0.2954</td>
<td>-1.667</td>
</tr>
<tr>
<td>COMDEL</td>
<td>-0.0186</td>
<td>-0.1123</td>
</tr>
<tr>
<td>ACWOT</td>
<td>1.000</td>
<td>10.41</td>
</tr>
</tbody>
</table>

---

\[ R^2 = 0.9125 \]
\[ DW = 1.05 \]

---

\[ R^2 = 0.9098 \]
\[ DW = 1.917 \]

---

\[ R^2 = 0.8088 \]
\[ DW = 1.5 \]

---

\[ R^2 = 0.8964 \]
\[ DW = 2.0 \]

---

\[ R^2 = 0.8521 \]
\[ DW = 1.597 \]

---

\[ R^2 = 0.8322 \]
\[ DW = 1.57 \]
Table 3

Production Model II

Medium Aggregated Delays

<table>
<thead>
<tr>
<th>Location</th>
<th>T Value</th>
<th>Coefficient</th>
<th>T Value</th>
<th>Coefficient</th>
<th>T Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>North 73</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>00.</td>
<td>-1.021</td>
<td>-1392.</td>
<td>-0.6773</td>
<td>-2287.</td>
<td>-3.448</td>
</tr>
<tr>
<td>1.06</td>
<td>-0.2430</td>
<td>12.76</td>
<td>0.4857</td>
<td>4.247</td>
<td>-0.4983</td>
</tr>
<tr>
<td>69.7</td>
<td>1.194</td>
<td>68.41</td>
<td>0.7221</td>
<td>75.29</td>
<td>2.244</td>
</tr>
<tr>
<td>58.5</td>
<td>6.240</td>
<td>-51.53</td>
<td>-0.4656</td>
<td>155.9</td>
<td>5.952</td>
</tr>
<tr>
<td>5.30</td>
<td>0.3376</td>
<td>29.97</td>
<td>0.4976</td>
<td>45.27</td>
<td>4.162</td>
</tr>
<tr>
<td>0.05</td>
<td>0.4176</td>
<td>-6.050</td>
<td>-0.1500</td>
<td>17.01</td>
<td>1.783</td>
</tr>
<tr>
<td>9.8</td>
<td>1.274</td>
<td>-20.28</td>
<td>-0.3184</td>
<td>-82.45</td>
<td>-2.970</td>
</tr>
<tr>
<td>619.</td>
<td>-2.103</td>
<td>46.97</td>
<td>0.5458</td>
<td>83.15</td>
<td>5.475</td>
</tr>
<tr>
<td>987</td>
<td>-1.549</td>
<td>-2644.</td>
<td>-4.719</td>
<td>-1638.</td>
<td>-2.705</td>
</tr>
<tr>
<td>129</td>
<td>-0.3559</td>
<td>-2.357</td>
<td>-0.9867</td>
<td>-1079.</td>
<td>-6.334</td>
</tr>
<tr>
<td>0.08661</td>
<td>-0.4794</td>
<td>-0.0604</td>
<td>-0.2107</td>
<td>-1729.</td>
<td>-6.016</td>
</tr>
<tr>
<td>0.909</td>
<td>6.984</td>
<td>0.0569</td>
<td>0.2032</td>
<td>0.143</td>
<td>20.39</td>
</tr>
<tr>
<td>South 73</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>74</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>96.</td>
<td>-1.385</td>
<td>-7947.</td>
<td>-1.664</td>
<td>-1.092</td>
<td>-0.4886</td>
</tr>
<tr>
<td>73</td>
<td>-0.5448</td>
<td>44.62</td>
<td>0.7431</td>
<td>128.6</td>
<td>1.539</td>
</tr>
<tr>
<td>2.4</td>
<td>0.9332</td>
<td>441.1</td>
<td>3.312</td>
<td>0.01382</td>
<td>0.05075</td>
</tr>
<tr>
<td>7.6</td>
<td>1.484</td>
<td>-103.6</td>
<td>-0.9025</td>
<td>-1751</td>
<td>-1.642</td>
</tr>
<tr>
<td>7.71</td>
<td>1.208</td>
<td>110.1</td>
<td>4.388</td>
<td>1.69</td>
<td>0.4777</td>
</tr>
<tr>
<td>3.32</td>
<td>0.2421</td>
<td>30.63</td>
<td>0.627</td>
<td>0.1283</td>
<td>0.4777</td>
</tr>
<tr>
<td>8.4</td>
<td>1.629</td>
<td>-115.0</td>
<td>-1.97</td>
<td>1.269</td>
<td>6.492</td>
</tr>
<tr>
<td>58.</td>
<td>-2.858</td>
<td>128.6</td>
<td>1.539</td>
<td></td>
<td></td>
</tr>
<tr>
<td>668</td>
<td>-0.4327</td>
<td>109.2</td>
<td>-0.4886</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1521</td>
<td>0.6656</td>
<td>0.01382</td>
<td>0.05075</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6526</td>
<td>1.639</td>
<td>0.1283</td>
<td>0.4777</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2175</td>
<td>0.9918</td>
<td>1.269</td>
<td>6.492</td>
<td></td>
<td></td>
</tr>
<tr>
<td>238</td>
<td>7.330</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

R^2 = 0.8088
Dw = 1.5

| South 74 |
|----------|---------|-------------|---------|
| 46.      | -1.385  | -7947.      | -1.664  |
| 73       | -0.5448 | 44.62       | 0.7431  |
| 2.4      | 0.9332  | 441.1       | 3.312   |
| 7.6      | 1.484   | -103.6      | -0.9025 |
| 7.71     | 1.208   | 110.1       | 4.388   |
| 3.32     | 0.2421  | 30.63       | 0.627   |
| 8.4      | 1.629   | -115.0      | -1.97   |
| 58.      | -2.858  | 128.6       | 1.539   |
| 668      | -0.4327 | 109.2       | -0.4886 |
| 1521     | 0.6656  | 0.01382     | 0.05075 |
| 6526     | 1.639   | 0.1283      | 0.4777  |
| 2175     | 0.9918  | 1.269       | 6.492   |
| 238      | 7.330   |             |         |

R^2 = 0.8322
Dw = 1.57
Table 4

Tests for Differences in Intercepts Across Sections

<table>
<thead>
<tr>
<th>Sections Included</th>
<th>E73</th>
<th>N73</th>
<th>S73</th>
<th>E74</th>
<th>N74</th>
<th>S74</th>
</tr>
</thead>
<tbody>
<tr>
<td>ENSEN, S74</td>
<td>B**</td>
<td>B</td>
<td>B</td>
<td>B</td>
<td>B</td>
<td>-48.9 (1.30)</td>
</tr>
<tr>
<td>E73,N73,S73</td>
<td>60.6 (0.35)</td>
<td>-597.5 (2.4)</td>
<td>B</td>
<td>-308.1 (1.8)</td>
<td>-121.3 (2.0)</td>
<td>-190.6 (0.99)</td>
</tr>
<tr>
<td>E74,N74,S74</td>
<td>9.1 (0.05)</td>
<td>-535.7 (2.4)</td>
<td>B</td>
<td>-103.4 (0.65)</td>
<td>-39.9 (1.98)</td>
<td></td>
</tr>
<tr>
<td>E73,E74</td>
<td>B</td>
<td>B</td>
<td></td>
<td>B</td>
<td>B</td>
<td>-115.4 (0.94)</td>
</tr>
<tr>
<td>N73,N74</td>
<td>B</td>
<td></td>
<td></td>
<td></td>
<td>372.2 (1.91)</td>
<td></td>
</tr>
<tr>
<td>S73,S74</td>
<td>B</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-163.9 (1.44)</td>
</tr>
</tbody>
</table>

* t-ratios are shown in parentheses.

** Sections denoted with a B in a given run were used as a base against which the remaining section intercepts were compared.
Table 5
Analysis of Variance Results

<table>
<thead>
<tr>
<th>Model II</th>
<th>DF</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>E73, N73, S73, E74, N74 vs. TOTAL</td>
<td>56/203</td>
<td>1.39</td>
</tr>
<tr>
<td>E73, N73, S73, E74, N74 vs. ENSEN</td>
<td>44/169</td>
<td>1.10</td>
</tr>
<tr>
<td>ENSEN, S74 vs. Total</td>
<td>12/247</td>
<td>3.28*</td>
</tr>
<tr>
<td>E73, N73, S73 vs. ENS 73</td>
<td>21/102</td>
<td>.99</td>
</tr>
<tr>
<td>ENS73, EN74 vs. ENSEN</td>
<td>12/201</td>
<td>1.29</td>
</tr>
<tr>
<td>E74, N74, S74 vs. ENS74</td>
<td>24/100</td>
<td>1.43</td>
</tr>
<tr>
<td>ENS73, ENS74 vs. Total</td>
<td>12/247</td>
<td>1.86**</td>
</tr>
<tr>
<td>E73, E74 vs. East 73 + 74</td>
<td>11/69</td>
<td>1.96**</td>
</tr>
<tr>
<td>N73, N74 vs. North 73 + 74</td>
<td>10/67</td>
<td>.85</td>
</tr>
<tr>
<td>S73, S74 vs. South 73 + 74</td>
<td>12/67</td>
<td>.93</td>
</tr>
</tbody>
</table>

* Significant at the 1% level.
** Significant at the 5% level.
Table 6

Tests for Differences Between Predicted and Actual Means

Control: South 73

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$\bar{y}$</td>
<td>4619</td>
</tr>
<tr>
<td>$\bar{y}_p$</td>
<td></td>
</tr>
<tr>
<td>$d = \bar{y}_e - \bar{y}_p$</td>
<td>97</td>
</tr>
<tr>
<td>$t$</td>
<td>.17</td>
</tr>
</tbody>
</table>

Control: Combined Nonexperimental Sections

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$\bar{y}$</td>
<td>4708</td>
</tr>
<tr>
<td>$\bar{y}_p$</td>
<td></td>
</tr>
<tr>
<td>$d = \bar{y}_e - \bar{y}_p$</td>
<td>8</td>
</tr>
<tr>
<td>$t$</td>
<td>.03</td>
</tr>
</tbody>
</table>
Appendix

DERIVATION OF THE TEST FOR
DIFFERENCES BETWEEN MEANS

The purpose of the test developed below is to compare mean weekly output between the experimental and control sections while controlling for differences in the values of all independent variables. The goal is to place as few restrictions as possible on the form and properties of the production function in the experimental section. Operationally the test involves substituting observed values of the independent variables from the experimental section into the estimated production function for the control section. The test then compares mean weekly production predicted from the function for the control sections to mean actual production from the experimental section.

We observe the values of the dependent variable generated in the experimental section. Call these \( y_e \), and let the mean of these be

\[
\bar{y}_e = \frac{1}{T} \sum_{t=1}^{T} y_{et} .
\]

We assume that \( y_{et} \) is normally distributed with unknown mean \( \mu_{e,t} \) and variance \( \sigma^2_e \). In general, \( \mu_{e,t} \) will be functionally dependent on values assumed by the vector of exogenous variables \( (X_t) \) in the experimental section at time \( t \), but it is not necessary to specify the form of that function for the test described below. Mean output in the experimental section is then normally distributed

\[
\bar{y}_e \sim N(\mu_{e}, \frac{\sigma^2_e}{T}).
\]
where

$$\mu_e = \frac{1}{T} \sum_{t=1}^{T} \mu_{e,t}$$

The vector of values \( (y_c) \) of the dependent variable in the control group is observed with a different matrix of exogenous variables \( Z \). Let \( Z \) have dimension \( N \times P \); there are \( N \) weekly observations on \( P \) variables in the control sections. Let \( X \) have dimension \( T \times P \) where there are \( T \) weekly observations on the vector of \( P \) independent variables, and the ordering of variables is the same as in \( Z \). The assumed model for the control group is linear:

(3) \( y_c = Z \beta + \varepsilon \)

where \( \varepsilon \sim N(0, \sigma_c^2 I_N) \) and \( Z \) and \( \varepsilon \) are independent.

If the matrix \( X \) were to occur in the control section, the model in (3) implies

(4) \( y_p = X \beta + u \)

where \( u \sim N(0, \sigma_c^2 I_T) \).

That is, \( y \) is the vector of observations on the dependent variable which would have occurred if the set of independent variables observed in the experimental section had also been observed in the control sections. The test procedure will be to predict \( y_p \) and to compare the mean of the prediction to \( y_e \), the mean of the observed values in the experimental section.
By taking the mean of the observations in (4) it follows that

\[
\bar{y}_p = \frac{1}{T} \sum_{t=1}^{T} y_{pt} = \bar{X} \beta + \bar{u}
\]

where \( \bar{u} \sim N(0, \frac{\sigma^2_u}{T} ) \) and \( \bar{X} \) is the vector of sample means of the variables in \( X \) (i.e., \( \bar{X}_t = \frac{1}{T} \sum_{t=1}^{T} x_{it} \)).

Then

\[
y_p \sim N(\mu_p, \frac{\sigma^2}{T} )
\]

where \( \mu_p = \bar{X}' \beta \).

Since \( \beta \) is unobservable, ordinary least squares estimators are obtained from (3).

\[
\hat{\beta} = (Z'Z)^{-1} Z' \bar{X}_t
\]

The mean of the predicted values from (4) is then

\[
\bar{y}_p = \bar{X}' \hat{\beta}.
\]

Since ordinary least squares is unbiased, it follows that

\[
E(\bar{y}_p) = \bar{X}' \beta = \mu_p.
\]

The variance of the predicted mean is

\[
\text{Var}(\bar{y}_p) = \text{E}[(\bar{y}_p - \mu_p)' (\bar{y}_p - \mu_p)]
\]

\[
= \text{E}[(\bar{X}' \hat{\beta} - \bar{X}' \beta)' (\bar{X}' \hat{\beta} - \bar{X}' \beta)]
\]

\[
= \frac{\sigma^2}{T} \bar{X}' (Z'Z)^{-1} \bar{X}.
\]
Then

\[(9) \quad \Delta y_p \sim N(\mu_p, \sigma_c^2 K)\]

where \(K = \bar{x}'(Z'Z)^{-1}\bar{x}\).

The objective is to test for the difference in means

\[(10) \quad d = \mu_p - \mu_e,\]

and the proposed test statistic is

\[(11) \quad \hat{d} = \frac{\Delta}{y_p - \bar{y}_e}.

The test statistic is unbiased.

\[E(d) = E(y_p - \bar{y}_e) = \mu_p - \mu_e = d\]

The variance of the test statistic is

\[\text{Var}(\hat{d}) = \text{Var}(\Delta y_p) + \text{Var}(\bar{y}_e) = K \sigma_c^2 + \frac{\sigma_e^2}{T}\]

where the assumption that \(u\) and \(e\) are independent has been utilized.

Thus

\[(12) \quad \hat{d} \sim N(\mu_p - \mu_e, K \sigma_c^2 + \frac{\sigma_e^2}{T})\]

The goal is to make a minimal set of assumptions about the experimental section. To avoid estimating the \(\mu_{e,t}\) which are functions of \(X_t\), we will assume that \(\sigma^2 = \sigma_e^2 = \sigma_c^2\). Then the variance can be estimated using the data from the control sections. The standard error of the regression in (7) is

\[(13) \quad \sigma_c^2 = \frac{1}{N-P} \sum_{t=1}^{T} (y_{ct} - \hat{y}_{ct})^2\]
From the normality of $\varepsilon$ it follows that

$$
(14) \quad \frac{(N-P)S^2}{\sigma^2} \sim \chi^2_{N-P}.
$$

The ratio of a unit normal and the square root of an independent chi-square divided by its degrees of freedom is distributed as the $t$.

Thus

$$
d - d \quad \frac{\sigma \sqrt{K + 1}}{\sqrt{(N - P)S^2}} \sim t_{N-P}.
$$

This reduces to the following

$$
(15) \quad \frac{d - d}{S \sqrt{K + 1}} \sim t_{N-P}
$$

This is the basis of the test results reported in the economic analysis of mean weekly production.
References


Goodman, Paul S., *Report: Rushton Quality of Work Project,*

DISTRIBUTION LIST

MANDATORY

Office of Naval Research
(Code 452)
800 N. Quincy Street
Arlington, Virginia 22217

Director
U.S. Naval Research Laboratory
Washington, D.C. 20390
ATTN: Technical Information Center

Defense Documentation Center
Building 5
Cameron Station
Alexandria, Virginia 22314

LIST 1

Library, Code 2029
U.S. Naval Research Laboratory
Washington, D.C. 20390

Science & Technology Division
Library of Congress
Washington, D.C. 20540

Psychologist
ONR Branch Office
495 Summer Street
Boston, Massachusetts 02210

LIST 2

Psychologist
ONR Branch Office
1030 E. Greet Street
Pasadena, California 91106

Research Psychologist
ONR Branch Office
536 S. Clark Street
Chicago, Illinois 60605

LIST 3

Principal Investigators

Dr. Earl A. Alluisi
Old Dominion University
Res. Foundation
Norfolk, Virginia 23508

Dr. Judith Daly
Decisions & Designs, Inc.
Suite 100
8400 Westpark Drive
McLean, Virginia 22101

Dr. James A. Bayton
Department of Psychology
Howard University
Washington, D.C. 20091

Dr. H. Russell Bernard
Department of Sociology & Anthropology
West Virginia University
Morgantown, West Virginia 26506

Dr. Arthur Blaives
Naval Training Equipment Center
Orlando, Florida 32813

Dr. Milton R. Blood
School of Business
Georgia Institute of Technology
Atlanta, Georgia 30332
Dr. David C. Bowers  
Institute for Social Research  
University of Michigan  
Ann Arbor, Michigan 48106

Dr. J. Richard Hackman  
Administrative Sciences  
Yale University  
56 Hillhouse Avenue  
New Haven, Connecticut 06520

Dr. John J. Collins  
Vice President  
Essex Corporation  
6305 Caminito Estrellado  
San Diego, California 92120

Dr. Leo A. Hazlewood  
CACI, Inc.  
1815 Fort Myer Dr.  
Arlington, Virginia 22209

Dr. Harry R. Day  
University City Science Center  
Center for Social Development  
3624 Science Center  
Philadelphia, Pennsylvania 19104

Dr. Edwin Hollander  
Department of Psychology  
State University of New York  
at Buffalo  
4230 Ridge Lea Rd.  
Buffalo, New York 14226

Dr. C. Brooklyn Derr  
Associate Professor, Code 55  
Naval Post Graduate School  
Monterey, California 93940

Dr. Daniel F. Huck  
General Research Corp.  
Westgate Research Park  
McLean, Virginia 22101

Dr. George T. Duncan  
Carnegie-Mellon University  
5000 Forbes Avenue  
Pittsburgh, Pennsylvania 15213

Dr. Charles L. Hulin  
Department of Psychology  
University of Illinois  
Champaign, Illinois 61820

Dr. Fred E. Fiedler  
Department of Psychology  
University of Washington  
Seattle, Washington 98105

Dr. Paul S. Goodman  
Graduate School of Industrial Admin.  
Carnegie-Mellon University  
Pittsburgh, Pennsylvania 15213

Dr. Edward E. Lawler  
Battelle Human Affairs Research Center  
4000 N.E. 41st Street  
P.O. Box 5395  
Seattle, Washington 98105

Dr. Rudi Klaus  
Syracuse University  
Public Administration Dept.  
Maxwell School  
Syracuse, New York 13210

Dr. Samuel L. Gaertner  
Department of Psychology  
University of Delaware  
2200 Wolf Hall  
Newark, Delaware 19711

Dr. Arie Y. Levin  
Duke University  
Duke Station  
Durham, North Carolina 27706

Dr. Paul E. Gayman  
Suite 200  
1055 Thomas Jefferson St., NW  
Washington, D.C. 20007

Dr. Morgan W. McCall, Jr.  
Center for Creative Leadership  
5000 Laurinda Dr.  
P.O. Box P-1  
Greensboro, North Carolina 27402
Dr. Terence R. Mitchell  
School of Business Administration  
University of Washington  
Seattle, Washington 98195

Dr. William H. Nobley  
College of Business Administration  
University of South Carolina  
Columbia, South Carolina 29208

Dr. Thomas D. Morris  
The Brookings Institution  
1775 Massachusetts Avenue NW  
Washington, D.C. 20036

Dr. James P. Murphy  
National Analysts  
A Division of Booz-Allen & Hamilton, Inc.  
400 Market Street  
Philadelphia, Pennsylvania 19106

Dr. Peter G. Nordlie  
Human Sciences Research, Inc.  
7710 Old Springhouse Rd.  
McLean, Virginia 22101

Dr. Herbert R. Northrup  
Industrial Research Unit  
University of Pennsylvania  
Philadelphia, Pennsylvania 19174

Dr. A. F. K. Organski  
3068 Institute for Social Research  
University of Michigan  
Ann Arbor, Michigan 48104

Dr. Paul Pedersen  
Society for Intercultural Education, Training and Research  
107 HUB, University of Pittsburgh  
Pittsburgh, Pennsylvania 15260

Dr. Chester M. Pierce  
Harvard University  
Nichols House  
Appian Way  
Cambridge, Massachusetts 02138

Dr. Manuel Ramirez  
Systems and Evaluations  
232 Swanton Blvd.  
Santa Cruz, California 95060

Dr. Irwin Sarason  
Department of Psychology  
University of Washington  
Seattle, Washington 98195

Dr. John Ruhe  
University of North Carolina  
Dept. of Business Admin.  
Charlotte, North Carolina 28223

Dr. Irwin Sarason  
Department of Psychology  
University of Washington  
Seattle, Washington 98195

Dr. S. B. Sells  
Texas Christian University  
Fort Worth, Texas 76129

Dr. R. Wallace Sinaiko  
Program Director  
Manpower Research & Advisory Services  
Smithsonian Institution  
801 N. Pitt St. - Suite 120  
Alexandria, Virginia 22314

Mrs. Alice I. Snyder  
Mental Health Clinic  
Naval Regional Medical Center  
Pearl Harbor  
FPO San Francisco 96610

Dr. Paul Spector  
American Institutes for Research  
Foxhall Square  
3301 New Mexico Avenue N.W.  
Washington, D.C. 20016

Dr. Richard Steers  
Graduate School of Management & Business  
University of Oregon  
Eugene, Oregon 97403

Dr. Lorna D. Szakay  
American Institutes for Research  
Foxhall Square  
3301 New Mexico Avenue N.W.  
Washington, D.C. 20016

Dr. Victor H. Vroom  
School of Organizational Management  
Yale University  
56 Hillhouse Avenue  
New Haven, Connecticut 06520
Dr. Paul Wall  
Division of Behavioral Science  
Tuskegee Institute  
Tuskegee, Alabama 36088

Dr. J. Wilkenfeld  
Department of Government & Politics  
College Park, Maryland 20742

Dr. Abraham R. Wagner  
Analytical Assessments Corp.  
357 South Robertson Blvd.  
Beverly Hills, California 90211

Dr. Davis E. Bobrow  
University of Maryland  
Department of Government and Politics  
College Park, Maryland 20742

---

LIST 4

---

Dr. A. L. Slafkosky  
Scientific Advisor  
Commandant of the Marine Corps (Code RD-I)  
Washington, D.C. 20380

HQ, USAF  
AFMPC/DPHYAR  
Randolph AFB, Texas 78148

Air University Library/  
LSE-8110  
Maxwell AFB, Alabama 36112

Army Research Institute  
Commonwealth Building  
1300 Wilson Blvd.  
Rosslyn, Virginia 22209

Dr. C. A. Patton, U.S.N.  
Director, Human Goals Dept.  
Code 70, Naval Training Ctr.  
Orlando, Florida 32813

Chief of Naval Personnel  
Assistant for Research  
Liaison (Pers-Or)  
Washington, D.C. 20370

Navy Personnel R & D Center  
Code 10  
San Diego, California 92152

Training Officer  
Human Resource Management Ctr.  
NTC, San Diego, California 92133

Human Resources Management Center Attachment  
Naval Support Activity  
c/o EP0 New York, NY 09521  
ATTN: TDC Nelson

Office in Charge (Code L5)  
Naval Aerospace Medical Research Lab.  
Naval Aerospace Medical Center  
Pensacola, Florida 32512

Journal Supplement Abstract Service  
APA  
1200 - 17th Street N.W.  
Washington, D.C. 20036

HumRRO (ATTN: Library)  
300 N. Washington Street  
Alexandria, Virginia 22314

Division Director for Social Science  
National Science Foundation  
1500 G Street N.W.  
Washington, D.C. 20550

APOSF (HL)  
1500 Wilson Blvd.  
Arlington, Virginia 22209
Chief, Psychological Research Branch  
U.S. Coast Guard (C-P-1/62)  
400 - 7th Street S.W.  
Washington, D.C. 20590

Bureau of Naval Personnel  
(Pers-6)  
Assistant Chief of Naval Personnel for Human Goals  
Washington, D.C. 20370

Scientific Director  
Naval Health Research Center  
San Diego, California 92152

Capt. Bruce G. Stone, USN  
(Code N-33)  
Director, Educ. & Training  
Research & Prog. Development

CNET Staff  
Naval Air Station  
Pensacola, Florida 32508

Director of Research  
HumRRO Division #4 (Infantry)  
P.O. Box 2086  
Fort Benning, Georgia 31905

Mr. Luigi Petrullo  
2431 N. Edgewood Street  
Arlington, Virginia 22207

Navy Materiel Command  
Employee Development Office  
Code SA-65  
Room 150 Jefferson Plaza  
Bldg. #2, 1429 Jeff Davis Highway  
Arlington, Virginia 20360

Personnel Research and Development Center  
United States Civil Service Commission  
Bureau of Policies and Standards  
Washington, D.C. 20415

Department of the Air Force  
Air Force Institute of Technology (AU)  
AFIT/SLGR (LT Col Umstot)  
Wright-Patterson Air Force Base, Ohio 45433